

EXPRESS MAIL CERTIFICATE

Date 12/14/01 Label No. 4767720048US

I hereby certify that, on the date indicated above I deposited this paper or fee with the U.S. Postal Service and that it was addressed for delivery to the Commissioner of Patents & Trademarks, Washington, DC 20231 by "Express Mail Post Office to Addressee" service.

Name (Print)

Signature

DOCKET NO.: 3343/11046-US1

5

10

## NAME SERVICE OBJECT CLUSTERING

### RELATED APPLICATIONS

The present patent application claims priority under 35. U.S.C. §119(e) from U.S. Provisional Patent Application Serial No. 60/255,651 Filed: December 14, 2000 filed on December 14, 2000, entitled Name Service Object Clustering, which is hereby incorporated herein by reference in its entirety.

### BACKGROUND OF THE INVENTION

#### 1. Field of the Invention

20

The present invention generally relates to the field of distributed object oriented computing and, more particularly, to a method for fault tolerance, load balance and "failover" of CORBA object servers via name service clustering.

#### 2. Description of the Related Art

25

Recently, Common Object Request Broker Architecture (CORBA) has emerged as the de facto standard architecture for distributed object computing. This distributed object

infrastructure enables organizations to build and deploy robust distributed object applications across the Internet and enterprise intranets.

Emerging application servers and Enterprise Java Bean (EJB) technology are being built on top of the pre-existing CORBA infrastructure. Remote Method Invocation (RMI) technology takes advantage of the capabilities of the Common Object Request Broker Architecture/Internet Inter Orb Protocol (CORBA/IIOP) infrastructure. As a result, fault tolerance, high availability and the load balancing capability of CORBA object servers become critically important due to the increased system capacity.

For example, a large number of requests to a CORBA object server can cause information bottlenecks at the server (i.e., a decrease in server capacity or a reduction in processing speed). As a result, a single CORBA object server can become the single-point of failure in a network. Moreover, since clients may be long lived, unavailability of services from a server is unacceptable.

In the past, these problems have been solved by replicating the CORBA servers and performing a load balance by using a predefined algorithm among the replicas to select a specific object server. Load balancing across the servers can be performed with a variety of techniques, such as via a Round Robin directory and a naming service or a special CORBA object server which dispatches requests from clients to servers. However, substantial coding, redesign and reconfiguration of the server, which results in increased costs, is unavoidable.

Another solution to the prior problems involves redesigning and replicating the CORBA server, and statically configuring the replica servers such that each server object contains one profile for each replica server object. Upon failure of a server, client requests "failover" to a replica server. However, this configuration is not scalable, i.e., it is not applicable to different servers located in a large network. Once the redesign, replication, and configuration of the server is completed and implemented, it cannot be changed.

A further solution to the above mentioned problems comprises redesigning the client such that it captures failures and redirects requests to another server. However, this approach not only requires client code changes, but also requires that the client possess prior knowledge related to the existence of all the servers in the network. It is therefore apparent  
5 there is a need for a method for reducing failures associated with selecting a server located in a distributed object oriented computing environment.

#### SUMMARY OF THE INVENTION

The present invention is directed to a method for fault tolerance, load balance  
10 and "failover" of CORBA object servers via name service clustering. Name service clustering, such as implemented in "Visibroker Naming Service" (Trademark, Borland Software Corp, Scotts Valley, CA), permits naming service load balancing over a set of object references contained within the same cluster such that loads are equitably distributed among servers. Each cluster contains its own unique object binding table which contains object references that each  
15 typically represent a single server. Load balancing is performed using a load balance algorithm (e.g., Round Robin). The specific algorithm which is used to perform load balancing is specified upon creation of each naming service cluster.

When a client machine invokes a cluster located under a particular context or specific directory, i.e., "resolves," a load balance is performed to return an object reference  
20 which was previously bound to the cluster. The client machine may then communicate with the server associated with the object reference which was selected. Failover is provided by dynamically adding cluster components to resolved object references upon resolving each cluster.

In accordance with the invention, a method for fault tolerance, load balance and  
25 failover of CORBA object servers comprises the steps of: invoking a cluster contained in a context; performing a load balance to select an object server located in the invoked cluster;

appending a cluster component to the invoked cluster to provide failover upon failure of the object server; forwarding a selected object reference to a client upon completion of the load balance; and communicating with a server associated with the selected object reference which was forwarded to the client.

5 In another embodiment of the invention a method for fault tolerance, load balance and failover of CORBA object servers comprises the steps of: setting a flag in a file to activate implicit clustering; invoking a cluster contained in a context having clusters; performing a load balance to select an object server located in the clusters; forwarding a selected object reference to a client upon completion of the load balance; and communicating  
10 with the server associated with the selected object reference which was forwarded to the client.

#### **BRIEF DESCRIPTION OF THE DRAWINGS**

The present invention may be more readily understood by one skilled in the art with reference being had to the following detailed description of the preferred embodiments  
15 thereof, taken in conjunction with the accompanying drawings wherein like elements are designated by identical reference numerals throughout the several views, and in which:

Fig. 1 is an illustration of a network arrangement of hardware components for implementing a method in accordance with a preferred embodiment of the invention;

Fig. 2 is an exemplary illustration of a naming graph comprising an object  
20 management group in a distributed computing system;

Fig. 3 is an illustration of a server replication operation in accordance with the preferred embodiment of the invention;

Fig. 4 is an illustration of a dynamic bind and unbind operation of cluster objects in accordance with the preferred embodiment of the invention;

Fig. 5 is a flow chart of a method for load balancing in accordance with the  
25 preferred embodiment of the invention; and

Fig. 6 is a flow chart of a method for dynamic object clustering in accordance with the preferred embodiment of the invention.

#### DESCRIPTION OF PREFERRED EMBODIMENTS

5 By way of overview and introduction, the invention is described in connection with a preferred embodiment, depicted in Fig. 3, in which load balancing is performed among servers 110 connected to the Internet 120, or other distributed computer network.

Generally, the Visibroker Naming Service™ permits the activation and deactivation of an implicit clustering feature and associates multiple object references with a  
10 single name. Activation of the implicit clustering feature may be accomplished using a flag located in a file, such as a configuration file, or the like. The clustering of object references in this manner becomes highly scalable because the object references may be dynamically added or removed from a cluster through cluster bind or unbind operations.

When implicit clustering is activated, object reference bindings using the same  
15 name are clustered together in the same cluster. A default load balance algorithm, such as Smart Round Robin, is used for implicit clustering. This algorithm may be changed at the discretion of the user.

If a call to a name within the clustered object references is invoked, the naming service load balances over the set of object references associated with the name and distributes  
20 the load accordingly. As a result, the need to recreate clusters is eliminated because the load balancing is performed among a common predetermined group of object reference binders.

Figure 1 shows a network arrangement by which the preferred embodiment can be implemented. The network 100 includes host servers 110, 110a...110n which provide content over the Internet 120 to a plurality of distributed users that access the host server  
25 through client machines 130, 130a...130n. The content provided by the host server 110,

110a...110n can be viewed by users through a web browser or other functionally equivalent software running at their respective client machines 130 (hereinafter, "browser").

The client machines or "client" 130 can assume a variety of forms, including a home computer, a personal digital assistant (PDA), a pager, an Internet compliant telephone, or other Internet compliant device. The client machines 130 and the host server 110 communicate with one another in a conventional manner over a communication link through a distributed computer network such as the Internet 120. In lieu of the Internet, communications can be through an Intranet or Extranet or between a host computer and a kiosk located nearby, as understood by those of skill in the art.

Figure 2 is an exemplary illustration of a naming graph comprising an object management group in a distributed computing system. The starting point for the construction of such a graph is the "Root Context" of the object management group. The Root context contains object name bindings. Such a repository of object name bindings is called a CosNaming Service, and permits a client to locate an object using a logical name which is bound within the repository.

Visibroker Naming Service™ is an implementation of OMG CosNaming Service. In addition, Visibroker Naming Service™ includes a feature called clustering which groups objects within a cluster to provide scalability and high availability to applications via load balancing and alternative server selection, i.e., "failover", upon server failures.

A Visibroker Naming Service™ cluster is a placeholder for a set of functionally exchangeable objects. A naming service cluster differs from the CosNaming context in that the CosNaming context is a directory that may contain sub-contexts, object references, and service clusters, while a service cluster may contain only object references. Each object reference shown in Fig. 2 will contain the server's Host Name, IP Address, Port No., and the like.

Figure 3 is an illustration of a server replication operation in accordance with the preferred embodiment of the invention. Visibroker Naming Service™ clustering permits

100495-12404

5 naming service load balancing over a set of object references contained within the same cluster such that loads are equitably distributed among servers. Each cluster contains its own unique object binding table which contains object references which each typically represent a single server. Load balancing is performed using a load balance algorithm (e.g., Round Robin). The specific algorithm which is used to perform load balancing in a given implementation is not critical to the invention, but is specified upon creation of each naming service cluster.

10 When a client invokes a cluster located under a particular context or specific directory (1), i.e., “resolves,” a load balance (2) is performed to return an object reference (3) which was previously bound to the cluster. The client may then communicate with the server associated with the object reference which was selected.

15 As shown in Fig. 3, failover is provided by dynamically adding cluster components to resolved object references upon resolving each cluster (4). In other words, upon resolving each cluster, a cluster component which is a part of an object reference of an object is created. Each cluster component provides to a client object resource broker (ORB) information which permits the client ORB to access the cluster object located in the name service to which this object is bound. Upon the occurrence of a failure, this information permits the client ORB to contact the cluster object to obtain another object in the same cluster to which the client ORB can then failover to. If an object reference fails during processing, the particular object reference binding associated with the specific object reference in the cluster becomes stale.

20

25 When a client invokes an invocation to a specific server, the connection to the server is accomplished using a process called binding. In order for the client to communicate with the server, it must bind with the server using the IP Address and Port No. contained in a specific object reference. Here, a binding is performed by using an object reference and performing the necessary steps to derive a concrete connection to a specific process that implements the functionality represented by that object reference. User participation in the

binding process is accomplished via a “binding interceptor” which is a Java class that the user may implement as desired. Each class comprises methods of differing relevancies which are used to indicate the binding status (i.e., Bind, Bind\_Succeeded, Bind\_Failed, and Exception\_Occurred). A “Bind” is performed by using an object reference and performing the

5 steps necessary to derive a concrete connection to a specific process which implements the functionality represented by that object reference. Here, a Bind\_Succeeded is an indication that a connection was achieved, a Bind\_Failed is an indication that a connection was not established and an Exception\_Occurred represents an indeterminate state between Bind\_Succeeded and Bind\_Failed.

10 Each time a client attempts to establish a connection with a server object, a bind method is called. In accordance with the invention, if a user is provided with the knowledge that bind interceptors may be used, the user will be provided with a class containing the most relevant methods, preferably the three most relevant methods. In the preferred embodiment, the most relevant methods are Bind, Bind\_Succeeded and Bind\_Failed. The user then writes

15 (specifies) the class such that the class subsequently contains the most relevant methods. When the connection to the server is being established, the user code is called (i.e., a bind interceptor is used). The bind interceptor provides a user with an option to select an alternative server in the instance where the system was operating adequately and then fails. For example, if a returned bind interceptor contains a Bind\_Failed method, the user is presented with an

20 opportunity to choose to connect to another server.

When an attempt to send a request to the stale server is made, the client associated with the stale object reference will transparently intercept the cluster component using the binding interceptor. The interceptor will then invoke the cluster’s select operation (i.e., its load balancing algorithm) to select and return to the client another object reference in

25 the same cluster. Communication is then established between the client and the server of the newly returned object reference. As a result, the failover to one of the object references in the

same cluster is transparently achieved. The failed object reference is thereafter marked as “suspect.” In an embodiment, the stale object reference is automatically removed from the cluster based on a user established preference.

In an aspect of the invention, the Visibroker Naming Service™ permits the activation and deactivation of an implicit clustering feature and associates multiple object references with a single name. Activation of the implicit clustering feature may be accomplished using a flag located in a file, such as a configuration file, or the like. The clustering of object references in this manner becomes highly scalable because the object references may be dynamically added or removed from a cluster through cluster bind or unbind operations.

Figure 4 is an illustration of a dynamic bind and unbind operation of cluster objects in accordance with the preferred embodiment of the invention. When implicit clustering is activated, object reference bindings using the same name are clustered together in the same cluster. A default load balance algorithm, such as Smart Round Robin, is used for implicit clustering. This algorithm may be changed at the discretion of the user.

When a call to a name within the clustered object references is invoked (1), the naming service load balances (2) over the set of object references associated with the name and distributes the load accordingly (4). As a result, the need to recreate clusters is eliminated because the load balancing is performed among a common group of object reference binders.

Fig. 5 is a flow chart of a method for load balancing in accordance with the preferred embodiment of the invention. In accordance with the preferred embodiment, the method is implemented when a client invokes a cluster located under a particular context or specific directory, as indicated in step 500. A load balance is performed to return an object reference which is bound to the cluster, as indicated in step 510. Load balancing is performed using a load balance algorithm (e.g., Round Robin). The specific algorithm which is used to perform load balancing is specified upon creation of each naming service cluster. An object

reference is forwarded to the cluster, as indicated in step 520. Next, cluster components are dynamically added to resolved object references subsequent to load balancing each cluster to provide failover, as indicated in step 530. Failover is provided in the manner discussed previously. Next, the client communicates with the server associated with the object reference  
5 which was selected and forwarded to the client, as indicated in step 540.

Figure 6 is a flow chart illustrating dynamic bind and unbind operations of cluster objects in accordance with the preferred embodiment of the invention. In accordance with this embodiment, the method is implemented when implicit clustering is activated, as indicated in step 600. Next, a cluster containing the clustered object reference bindings using  
10 the same name is invoked, as indicated in step 610. Load balancing is performed on the clustered object references using a default load balance algorithm, as indicated in step 620. This algorithm may be changed at the discretion of the user, e.g., from a Round Robin algorithm to a different algorithm. A selected object reference based on the load balance is returned upon completion of load balancing, as indicated in step 630. Next, the client  
15 communicates with the server associated with the object reference which was selected and forward to the client, as indicated in step 640.

Methods in accordance with the invention permit transparent load balancing among multiple object references in the same cluster to be achieved. In addition, transparent failover among object references in the same cluster is provided. Method according to the  
20 invention also provide a dynamic addition of new object references into a cluster via binding, along with the dynamic removal of old object references from a cluster via unbinding. Moreover, automatic detection of stale object references in a cluster and clean up based on the preference of a client is achieved. Further, such methods provide a consistent application of the load balancing algorithms per cluster. New load balancing algorithms may be up-loaded  
25 to the system as required. The method eliminates the need to change the client code or server code such that manual configurations are not required.

While the invention has been particularly shown and described with reference to a preferred embodiment thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the spirit and scope of the invention.